Mathematics 195A—Honors Seminar
Winter 2005

Bernard Russo

March 6, 2005

For all aspects of the course, consult

`http://math.uci.edu/~brusso/.`

1 January 5,2005—The Prime Number Theorem I

Our first objective is to explore an elementary proof of the prime number theorem, following [6].

The prime number theorem was first conjectured by Gauss and Legendre at the end of the 18th century. It was proved, using complex analysis, at the end of the 19th century by de la Vallée Poussin and Hadamard. In the middle of the 19th century, significant tools were developed by Chebyshev and Riemann. An elementary proof, that is, not using complex analysis, was discovered in the middle of the 20th century by Erdös and Selberg.

Let $\pi(x) := \sum_{p\le x} 1$ be the number of primes less than the positive number x and let $Li(x) := \int_2^x \frac{dt}{\log t}$ be the "log integral" function. Legendre conjectured that $\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1$ and Gauss conjectured that $\lim_{x\to\infty} \frac{\pi(x)}{Li(x)} = 1$.

For a history of the prime number theorem and the Riemann hypothesis, see the NY Times bestseller [4], and my Freshman seminar (University Studies 3, Winter 2005).

2 January 7,2005—The Prime Number Theorem II

By the fundamental theorem of arithmetic, each positive integer $n \ne 1$ has the form $n = p_1^{k_1}\cdots p_m^{k_m}$, where $m \ge 1, p_1,\ldots,p_m$ are distinct primes and $1 \le k_j$, so that $\log n = \sum_{j=1}^m k_j \log p_j$. Define the **von Mangoldt symbol** (1895) by $\Lambda(n) := \log p$ if n is a positive integral power of the prime p, and $\Lambda(n) = 0$ if n is divisible by the square of some prime.

Exercise 1 $\log n = \sum_{j|n} \Lambda(j)$.

Define functions ψ, θ, T as follows: $\psi(x) := \sum_{j\le x} \Lambda(j)$; $\theta(x) := \sum_{p\le x} \log p$; $T(x) := \sum_{n\le x} \log n$. Then (for example) $\psi(x) = \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \cdots$. The prime number theorem is the assertion $\lim_{x\to\infty} \pi(x)/(x/\log x) = 1$, which is equivalent to $\lim_{x\to\infty} \psi(x)/x = 1$. This equivalence, as well as the proof of the latter are the objectives of [6] and the first week or two of our class.

For a function $F(x)$ defined for $x > 1$, define a transform $F \mapsto G$ by $G(x) = \sum_{n\le x} F(x/n) = F(x) + F(x/2) + F(x/3) + \cdots + F(x/[x])$.

Proposition 2.1 *$F(x) = \sum_{k\le x} \mu(k) G(x/k)$ where μ is defined as follows: $\mu(1) = 1$, $\mu(n) = (-1)^m$ if n is the product of m distinct primes, and $\mu(n) = 0$ otherwise (that is, $\mu(n) = 0$ if n is divisible by the square of some prime).*

Proof. See [6, pp.228-230].

Applying this *Möbius inversion formula* to $T(x) = \sum_{i \le x} \psi(x/i)$, you get $\psi(x) = \sum_{k \le x} \mu(k) T(x/k)$ which can be rewritten (see [6, p.230]) as

$$\sum_{n \le x} \Lambda(n) = \sum_{n \le x} \sum_{k|n} \mu(k) \log(n/k).$$

Exercise 2 $\Lambda(n) = \sum_{k|n} \mu(k) \log(n/k)$

3 January 10,2005—The Prime Number Theorem III

The following five lemmas are the objective of today's class. In the class we only proved the second and fourth (Lemmas 3.2 and 3.4 in [6]). For the proof of the others, please consult [6].

Lemma 3.1 (**[6, Lemma 3.1]**) *If $f'(t)$ is continuous for $t \ge 1$, and $C(u) := \sum_{n \le u} c_n$ for some sequence of numbers $\{c_n : n \ge 1\}$, then*

$$\sum_{n \le x} c_n f(n) = f(x)C(x) - \int_1^x f'(t)C(t)\, dt,$$

and

$$\sum_{n \le x} f(n) = \int_1^x f(t)\, dt + \int_1^x (t - [t]) f'(t)\, dt + f(1) - (x - [x]) f(x).$$

The notation $f(x) = O(g(x))$ $(x \to \infty)$ for a function f and a non-negative function g means that there are constants K_1 and K_2 such that $|f(x)| \le K_1 g(x)$ for all $x \ge K_2$.

Exercise 3 *Put $f(t) = \log t$ in Lemma 3.1 to get $T(x) = x \log x - x + O(\log x)$.*

Lemma 3.2 (**[6, Lemma 3.2]**) *$\psi(x) < (3/2)x$ for large x.*

Lemma 3.3 (**[6, Lemma 3.3]**) $\sum_{n \le x} \Lambda(n)/n = \log x + O(1)$

Lemma 3.4 (**[6, Lemma 3.4]**) $\psi(x) = \pi(x) \log x + O(x \log\log x / \log x)$

Exercise 4 *Show that the inequality*

$$\pi(x) \log x - \frac{4x \log\log x}{\log x} \frac{x}{\log x} \le \psi(x) \le \pi(x) \log x + \frac{x^{1/2} (\log x)^2}{2 \log 2}$$

proved in [6, pp.233-234] implies the assertion of Lemma 3.4.

Exercise 5 *Use Lemma 3.4 to show that* $\lim_{x\to\infty} \psi(x)/x = 1$ *if and only if* $\lim_{x\to\infty} \pi(x)/(x/\log x) = 1$

Lemma 3.5 (**[6, Lemma 3.5]**) $\sum_{n \le x} 1/n = \log x + \gamma + O(1/x)$

4 January 14,2005—The Prime Number Theorem IV

Let's apply Möbius inversion to $F = \psi \mapsto G = T$ with the following strategy. Start with $\tilde{F}$ simpler than F and try to make $\tilde{G}$ close to T. Möbius inversion then gives you

$$\psi(x) - \tilde{F}(x) = \sum_{k \le x} \mu(k) (T(x/k) - \tilde{G}(x/k)) \quad \text{(equation (4.1))}.$$

Since the desired goal is $\psi(x)/x \to 1$, we initially choose $\tilde{F}(x) = F_0(x) = x$ which results (using Lemma 3.5) in $G_0(x) = x \log x + x\gamma + O(1)$. Using $T(x) = x \log x - x + O(\log x)$, this results in $T(x) - G_0(x) =$

$-x(1+\gamma)+O(\log x)$. This is not good enough for our purposes, so the next guess is $\tilde{F}(x) = F_1(x) = x - C$. This results in $G_1(x) = x\log x - (C-\gamma)x + O(1)$. Then choosing $C = 1+\gamma$, you get

$$T(x) - G_1(x) = O(\log x) \quad \text{(equation (4.2))}.$$

By (4.1) with $\tilde{F}(x) = x - C$

$$\psi(x) - x + C = \sum_{k\le x} \mu(k)(T(x/k) - G_1(x/k)) \quad \text{(equation (4.3))},$$

and $T(x/k) - G_1(x/k) = O(\log(x/k))$.

Even if we replace the $O(\log x)$ implicit in (4.3) by $O(x^{1/2})$, we can still derive the (known fact)

$$\psi(x) = O(x) \quad \text{(equation (4.4))},$$

as shown by equation (4.6). This suggests the Tatuzawa-Iseki identity (at least to the author Norman Levinson!), which states

$$F(x)\log x + \sum_{n\le x} F(\frac{x}{n})\Lambda(n) = \sum_{k\le x} \mu(k)\log\frac{x}{k}G(\frac{x}{k}) \quad \text{(equation (4.9))},$$

and which leads easily to the inequality of Selberg, which states

$$(\psi(x) - x)\log x + \sum_{n\le x}(\psi(\frac{x}{n}) - \frac{x}{n})\Lambda(n) = O(x) \quad \text{(equation (4.10))}.$$

Exercise 6 *Prove the following, which was used in the proof of (4.10).*

$$\sum_{k\le x}\mu(k)\log(\frac{x}{k})(T(\frac{x}{k}) - G_1(\frac{x}{k})) = O(x).$$

There are eight lemmas in [6, section 5] which constitute the proof of PNT. We now state and prove the first one.

Define $R(x) = \psi(x) - x$ for $x \ge 2$ and $R(x) = 0$ for $0 < x < 2$. Then PNT is obviously equivalent to $R(x)/x \to 0$. Define $S(y) = \int_2^y R(x)/x\,dx$ for $y \ge 2$ and $S(y) = 0$ for $0 < y < 2$. Later, we will show that if $S(y)/y \to 0$ then $S(x)/x \to 0$, whence PNT. Of course, we have also to prove $S(y)/y \to 0$!

Lemma 4.1 ([**6, Lemma 5.1**]) *There is a constant c such that*

(equation (5.5)) $|S(y)| \le cy$ *for* $y \ge 2$

(equation (5.6)) $|S(y_2) - S(y_1)| \le c|y_2 - y_1|$

(equation (5.7)) $S(y)\log y + \sum_{j\le y}\Lambda(j)S(y/j) = O(y)$

5 January 17,2005—Holiday

6 January 21,2005—The Prime Number Theorem V

By using Lemma 3.3, equation (4.10) above can be rewritten as

$$\psi(x)\log x + \sum_{n\le x}\Lambda(n)\psi(\frac{x}{n}) = 2x\log x + O(x). \quad \text{(equation (4.11))}$$

Using Lemma 3.1 with $c_n = \Lambda(n)$ and $f(t) = \log t$ and Lemma 3.2 results in

$$\sum_{n\le x}\Lambda(n)\log n = \psi(x)\log x + O(x). \quad \text{(equation (4.12))}$$

Also

$$\sum_{j\le x}\Lambda(j)\psi(\frac{x}{j}) = \sum_{j\le x}\Lambda(j)\sum_{k\le x/j}\Lambda(k) = \sum_{jk\le x}\Lambda(j)\Lambda(k). \quad \text{(equation (4.13))}$$

Thus, if we define $\Lambda_2(n) := \Lambda(n)\log n + \sum_{jk=n}\Lambda(j)\Lambda(k)$ and plug (4.12) and (4.13) into (4.11), we get $\sum_{n\le x}\Lambda_2(n) = 2x\log x + O(x)$. From Exercise 3 you get $\sum_{n\le x}\log n = x\log x + O(x)$. Finally, if we define $Q(n) := \sum_{k\le n}(\Lambda_2(k) - 2\log k)$, then

$$Q(n) = O(n) \quad \text{(equation (4.15))}$$

for $n \ge 2$ while $Q(1) = 0$.

Lemma 6.1 (**[6, Lemma 5.2]**) *There is a constant K_1 such that*

$$\log^2 y|S(y)| \le \sum_{m\le y}\Lambda_2(m)|S(\frac{y}{m})| + K_1 y\log y. \quad \text{(equation (5.13))}$$

7 The Prime Number Theorem VI (not done in class)

Lemma 7.1 (**[6, Lemma 5.3]**) *There is a constant K_2 such that*

$$\log^2 y|S(y)| \le 2\sum_{m\le y}\log m|S(\frac{y}{m})| + K_2 y\log y. \quad \text{(equation (5.14))}$$

Lemma 7.2 (**[6, Lemma 5.4]**) *There is a constant K_4 such that*

$$\log^2 y|S(y)| \le 2\int_2^y |S(\frac{y}{u})\log u\,du + K_4 y\log y. \quad \text{(equation (5.16))}$$

In (5.16), let $v = \log(y/u)$ and $x = \log y$. Then

$$x^2|S(e^x)| \le 2\int_0^{x-\log 2}|S(e^v)|(x-v)e^{x-v}\,dv + K_4 x e^x. \quad \text{(equation (5.18))}$$

Set $W(x) := e^{-x}S(e^x)$. Then (5.18) becomes

$$|W(x)| \le \frac{2}{x^2}\int_0^x (x-v)|W(v)|\,dv + \frac{K_4}{x}. \quad \text{(equation (5.20))}$$

Lemma 7.3 (**[6, Lemma 5.5]**)

$$\alpha := \limsup_{x\to\infty}|W(x)| \le \min\{1, \gamma := \limsup_{x\to\infty}\frac{1}{x}\int_0^x |W(\xi)|\,d\xi\} \quad \text{(equation (5.22))}$$

NOTE: PNT will follow from the assertion $\alpha = 0$.

Lemma 7.4 (**[6, Lemma 5.6]**) *If $k := 2c$, then*

$$|\,|W(x_2)| - |W(x_1)|\,| \le |W(x_2) - W(x_1)| \le k|x_2 - x_1| \quad \text{(equations (5.26) and (5.27))}$$

Lemma 7.5 (**[6, Lemma 5.7]**) *If $W(v) \ne 0$ for $v_1 < v < v_2$, then $\exists M > 0$ such that*

$$\int_{v_1}^{v_2}|W(v)|\,dv \le M \quad \text{(equation (5.28))}$$

Lemma 7.6 (**[6, Lemma 5.8]**) *If a function W satisfies (5.22), (5.27), and (5.28), then $\alpha = 0$.*

Discussion: The proofs of Lemmas 7.1-7.5, as well as the proof that PNT follows from $\alpha = 0$ are easy to follow from [6]. Lemma 7.6 is another matter.

Exercise 7 *Give an understandable proof of Lemma 7.6.*

8 January 24,2005—Continued Fractions I

Consider the following problem: given positive integers a, b, c, obtain solutions of the Diophantine equation $ax \pm by = c$ (equation (4.1)). It is enough to consider the case with the plus sign, and we can assume that a and b have no common factor.

Write $a/b = \beta_0 + 1/r_1$, where $\beta_0 = [a/b]$ and $1 < r_1 \leq \infty$. (The meaning here of "$r_1 = \infty$" is that a/b is an integer, so the construction ends.) If "$r_1 \neq \infty$", write $r_1 = \beta_1 + 1/r_2$, where $\beta_1 = [r_1]$ and $1 < r_2 \leq \infty$. At this point we have

$$\frac{a}{b} = \beta_0 + \frac{1}{\beta_1 + \frac{1}{r_2}} \quad \left(\text{or } \beta_0 + \frac{1}{\beta_1} \text{ if } r_2 = \infty \right)$$

Continue this construction to obtain $r_n = \beta_n + 1/r_{n+1}$, where $\beta_n = [r_n]$ and $1 < r_n \leq \infty$. This construction ends in finite sequences $\beta_0, \beta_1 \ldots, \beta_n$ and $r_1, r_2, \ldots, r_n$ if some $r_{n+1} = \infty$; otherwise it is an infinite process generating infinite sequences $\beta_0, \beta_1 \ldots$ and $r_1, r_2, \ldots$. Therefore we have

$$\frac{a}{b} = \beta_0 + \frac{1}{\beta_1 + \frac{1}{\beta_2 + \frac{1}{\beta_3 + \cdots}}} \quad \left(\text{or } \beta_0 + \frac{1}{\beta_1 + \frac{1}{\beta_2 + \frac{1}{\beta_3}}} \text{ if } r_4 \text{ (for example) } = \infty \right)$$

This suggests considering expressions of the form

$$\beta_0 + \frac{\alpha_1}{\beta_1 + \frac{\alpha_2}{\beta_2 + \frac{\alpha_3}{\beta_3 + \cdots}}}$$

where $\{\alpha_i\}_{i \geq 1}$ and $\{\beta_i\}_{i \geq 0}$ are sequences of real numbers. For sanity's sake, we shall denote such an expression (which could be finite or infinite) by

$$\beta_0 + \frac{\alpha_1}{\beta_1 +} \frac{\alpha_2}{\beta_2 +} \frac{\alpha_3}{\beta_3 +} \cdots. \tag{1}$$

Given the continued fraction (1), consider the *convergents*

$$Q_n = Q_n(\beta_0, \alpha_1, \beta_1, \cdots, \alpha_n, \beta_n) = \beta_0 + \frac{\alpha_1}{\beta_1 +} \frac{\alpha_2}{\beta_2 +} \frac{\alpha_3}{\beta_3 +} \cdots \frac{\alpha_n}{\beta_n}.$$

The continued fraction (1) *converges* if $\lim_n Q_n$ exists, and we write $\beta_0 + \frac{\alpha_1}{\beta_1 +} \frac{\alpha_2}{\beta_2 +} \cdots = \lim_n Q_n$.

Given the two sequences $\{\alpha_i\}_{i \geq 1}$ and $\{\beta_i\}_{i \geq 0}$, consider the two three-term recurrence sequences

$$R_{-1} = 1, \quad R_0 = \beta_0, \quad \text{and for } n \geq 1, \quad R_n = \beta_n R_{n-1} + \alpha_n R_{n-2},$$

and

$$S_{-1} = 0, \quad S_0 = 1, \quad \text{and for } n \geq 1, \quad S_n = \beta_n S_{n-1} + \alpha_n S_{n-2}.$$

It is important to note that $R_n = R_n(\beta_0, \alpha_1, \beta_1, \cdots, \alpha_n, \beta_n)$ and $S_n = S_n(\alpha_1, \beta_1, \cdots, \alpha_n, \beta_n)$.

Proposition 8.1 J. Wallis 1655 ([**3, section 4.1**])
For the continued fraction (1), $Q_n = R_n/S_n$ *for every* n.

Proposition 8.2 ([**3, section 4.1**])
$R_n S_{n-1} - R_{n-1} S_n = (-1)^{n+1} \alpha_1 \cdots \alpha_n$ *for every* $n \geq 1$.

Exercise 8 *The proof of Proposition 8.2 was given under the assumption that* $\alpha_1 \cdots \alpha_n \neq 0$. *What is the proof in case some* α*s are zero?*

9 January 28,2005—Continued Fractions II

Theorem 9.1 ([**3, Theorem 4.8**]) *For each real number* γ, *there is a unique continued fraction with value* γ *of the form*

(i) (γ irrational) $\gamma = \beta_0 + \frac{1}{\beta_1 +} \frac{1}{\beta_2 +} \cdots$ *with* $\beta_0 \in \mathbf{Z}$ *and* $\{\beta_i\}_{i \geq 1}$ *positive integers.*

(ii) (γ rational) $\gamma = \beta_0 + \frac{1}{\beta_1 +} \frac{1}{\beta_2 +} \cdots \frac{1}{\beta_n}$ *with* $\beta_0 \in \mathbf{Z}$ *and* $\{\beta_i\}_{1 \leq i \leq n}$ *positive integers*

Exercise 9 *Prove the uniqueness part of Theorem 9.1.*

10 January 31,2005—Continued Fractions III

10.1 Application to a Diophantine Equation

We return to the Diophantine equation $ax \pm by = c$ (equation (4.1)). We know that $a/b = Q_n = R_n/S_n$ where $R_j = \beta_j R_{j-1} + R_{j-2}$ and $S_j = \beta_j S_{j-1} + S_{j-2}$ for $1 \leq j \leq n$ and the initial conditions are $R_{-1} = 1, R_0 = \beta_0, S_{-1} = 0, S_0 = 1$. Since $R_n S_{n-1} - R_{n-1} S_n = (-1)^{n+1}$ we have $(R_n, S_n) = 1$, and since $(a, b) = 1$, we have $a = R_n$ and $b = S_n$. It follows that for every t, $x := bt + (-1)^{n+1} c S_{n-1}$ and $y := -at - (-1)^{n+1} c R_{n-1}$ are solutions of $ax + by = c$ which are integers if t is an integer.

10.2 Suggestions for projects on continued fractions

Quadratic irrationals and continued fractions References: two papers of Lewittes ([7],[8]) and the book of Ono ([9]).

Applications of continued fractions Chapter 4 of the book by Rockett and Szüsz, [10].

Continued fractions and orthogonal polynomials Searching the AMS website (MathSciNet) using the key words "continued fractions" and "orthogonal polynomials" leads to 191 entries!

10.3 Regular continued fractions

(This subsection is from [10, p. 3-4].)

Another notation for the continued fraction (1) with all the $\alpha_j = 1$ is $[\beta_0; \beta_1, \ldots, \beta_n, \ldots]$. A *regular continued fraction* is one for which β_0 is an integer and β_k is a positive integer for $k \geq 1$. In such a case, we have $(R_k, S_k) = (R_k, R_{k+1}) = (S_k, S_{k+1}) = 1$ and $t = \lim_{k \to \infty} R_k/S_k$ exists.

Exercise 10 *Show that R_k/S_k approximates t alternatively from above and below.*

If $t = [a_0; a_1, \ldots, a_n]$ is a regular continued fraction, then t is rational. If $a_n > 1$, then $t = [a_0; a_1, \ldots, a_n - 1, 1]$; if $a_n = 1$, then $t = [a_0; a_1, \ldots, a_{n-1} + 1]$. You could have uniqueness by insisting that $a_n \geq 2$, but we won't do this. Finally, if $t = [a_0; a_1, \ldots, a_n, \ldots]$ doesn't terminate, then t is irrational.

11 February 4,2005—No class

12 February 7,2005—Braid Group I

For this topic, we are following [5].

Braids can be made of several types of material (e.g., rope, hair, dough), can have cultural significance (e.g., Ukrainian bread, Mexican belts), and can occur in nature (e.g., rings of Saturn, DNA, periodic orbits).

The definition of a braid must use mathematical concepts and ideas. A braid is a geometric object, and the material it is made of is irrelevant. Algebra is used to study properties of braids. Braids were developed first by Emil Artin in two papers (1925—a geometric approach, in German [1]; 1947—an algebraic approach, in English [2])

An *n-braid* consists of the unit cube $\mathbf{D}$ in $\mathbf{R}^3$, n points $A_1, \ldots, A_n$ on the top of the cube, n points $B_1, \ldots, B_n$ on the bottom and n polygonal segments $d_1, \ldots, d_n$ (called *braid strings* and drawn as smooth arcs) which satisfy the following conditions

- $d_1, \ldots, d_n$ are pairwise disjoint
- Each d_i connects some A_j to some B_k
- Each horizontal place $E_s = \{(x, y, z) : 0 \leq x, y \leq 1, z = s\}$, with $0 \leq s \leq 1$ meets each d_j in exactly one point.

The set of all n-braids is denoted $\mathcal{B}_n$. Two braids are said to be *equivalent* if one can be obtained from the other with a finite sequence of *elementary moves.* An *elementary move* on a braid is the process of replacing a segment of one string d, by two segments which together with the original segment forms a triangle which doesn't intersect any other string and intersects d only in this segment. (The inverse process is also considered to be an elementary move). This is an equivalence relation $\beta \sim \beta'$, and $\mathbf{B}_n = \mathcal{B}_n/\sim$ denotes the set of all equivalence classes.

13 February 11,2005—Braid Group II

Braids are visualized by means of the braid projection $p : \mathbf{D} \to \mathbf{D}$, $p(x, y, z) = (0, y, z)$. By performing some elementary moves on a braid β, we assume the curves $p(d_i)$ satisfy

- $p(\beta)$ has only a finite numbe of intersection points
- If Q is such an intersection point (called a *double point*), then $p^{-1}(Q) \cap \beta$ has exactly two points
- A vertex (obvious definition) of β is never mapped by p onto a double point of $p(\beta)$.

At this point, $p(\beta)$ represents β except at double points. To indicate which string is in front of the other, the projection diagram (but not the string!) which is behind the other one is cut.

Non-equivalence of braids can be shown by use of *invariants*, that is, functions $f : \mathcal{B}_n \to$ some algebraic structure such that $\beta \sim \beta' \Rightarrow f(\beta) = f(\beta')$. Simple examples of invariants are: $f(\beta)$= the number of strings of β; and

$$f(\beta) = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ j(1) & j(2) & j(3) & \cdots & j(n) \end{pmatrix}$$

the braid permutation, where d_i connects A_i to $B_{j(i)}$.

Theorem 13.1 ([**5, Theorem 1.5,p.15**]) *$\mathbf{B}_n$ is a group, under $[\beta][\beta'] = [\beta\beta']$, where $[\beta]$ is the equivalence class of $\beta \in \mathcal{B}_n$ and $\beta\beta'$ is the multiplication of braids, obtained by putting the projection diagram of β on top of the projection diagram of β' and removing the horizontal line through the points of connection.*

14 February 14,2005—Braid Group III

You can partition any braid diagram by horizontal lines such that between two consecutive lines, only two strings are braided with a solitary double point and the other strings remain vertical. This immediately leads to the conclusion that the braid group $\mathbf{B}_n$ is generated by $n-1$ elements $\sigma_1, \ldots, \sigma_{n-1}$. These generators satisfy two types of relations, $\sigma_i\sigma_j = \sigma_j\sigma_i,\ 1 \le i < j \le n-1,\ j - i \ge 2$, and $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1},\ 1 \le i \le n-2$, which leads to a (so-called) *presentation* of the group $\mathbf{B}_n$.

Theorem 14.1 ([**5, Theorem 2.2,p.18**])

$$\mathbf{B}_n = \langle \sigma_1, \ldots, \sigma_{n-1} | \sigma_i\sigma_j = \sigma_j\sigma_i,\ \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \rangle$$

For the present, we shall take the meaning of "presentation" to be that the group is specified by a set of generators and a set of relations satisfied by those generators. We do not at this time address the precise meaning of this, which is explained in the appendix of [5].

15 February 18,2005—Braid Group IV and V

15.1 Free Groups

Let $S = \{x_1, \ldots, x_n\}$ be a set and let $S^{-1} = \{x_1^{-1}, \ldots, x_n^{-1}\}$ be another set with the same number of elements (n is supposed finite, but the same reasoning will apply to a set of any cardinality). A *word* in $S \cup S^{-1}$ is an expression $W = x_{i_1}^{\epsilon_1} x_{i_2}^{\epsilon_2} \cdots x_{i_k}^{\epsilon_k}$, where $1 \le i_1, \ldots, i_k \le n,\ \epsilon_i = \pm 1$. Let $\mathcal{W}$ be the set of all such words, together with the empty word, denoted by 1 and define the product of words by juxtaposition: $W_1W_2 = x_{i_1}^{\epsilon_1}x_{i_2}^{\epsilon_2}\cdots x_{i_p}^{\epsilon_p}y_{i_1}^{\eta_1}y_{i_2}^{\eta_2}\cdots y_{i_q}^{\epsilon_q}$ and $1W_1 = W_1 1 = W_1$ if $W_1 = x_{i_1}^{\epsilon_1}x_{i_2}^{\epsilon_2}\cdots x_{i_p}^{\epsilon_p}$ and $W_2 = y_{i_1}^{\eta_1}y_{i_2}^{\eta_2}\cdots y_{i_q}^{\epsilon_q}$. Clearly, $\mathcal{W}$ is an associative semigroup with identity.

Define two words to be equivalent if you can get from one to the other by a finite sequence of "insertions" and "deletions" of terms of the form $x_p^{\epsilon_p}x_p^{-\epsilon_p}$.

Theorem 15.1 (**[5, Theorem 3.1,p.233]**) *The set $\tilde{\mathcal{W}}$ of equivalence classes is a group under $[W_1][W_2] = [W_1W_2]$ and $[W]^{-1} = [x_{i_p}^{-\epsilon_p}\cdots x_{i_2}^{-\epsilon_2}x_{i_1}^{-\epsilon_1}]$ if $W = x_{i_1}^{\epsilon_1}x_{i_2}^{\epsilon_2}\cdots x_{i_p}^{\epsilon_p}$.*

$\tilde{\mathcal{W}}$ is said to be a free group of rank n and is denoted by $F\langle x_1,\ldots,x_n\rangle$.

Theorem 15.2 (**[5, Theorem 3.2,p.233]**) *Two free groups of the same rank are isomorphic.*

The free group $F = F\langle x_1,\ldots,x_n\rangle$ has the following universal property. Let G be any group with n generators $g_1,\ldots,g_n$. Then the map $f : x_i \mapsto g_i$ $(1 \le i \le n)$ extends to a homomorphism $\hat{f}$ of F onto G, given by $\hat{f}(x_{i_1}^{\epsilon_1}x_{i_2}^{\epsilon_2}\cdots x_{i_p}^{\epsilon_p}) = g_{i_1}^{\epsilon_1}g_{i_2}^{\epsilon_2}\cdots g_{i_p}^{\epsilon_p}$.

15.2 The word problem

Given a group G represented as $G = \langle x_1,\ldots,x_n|R_1 = 1,\ldots,R_m = 1\rangle$, the *word problem* for G is to find a "reasonably practical" method that will be able to decide whether or not two arbitrary words (=elements of G) g_1 and g_2 are equal; equivalently, given $g \in G$, when is $g = 1$?

Theorem 15.3 (**[5, Theorem 5.1,p.239]**) *The word problem is solvable for the free groups.*

Proof. A word $g = x_{i_1}^{\epsilon_1}x_{i_2}^{\epsilon_2}\cdots x_{i_p}^{\epsilon_p}$ is equal to 1 if it is either the empty word or if we can eliminate each $x_{i_j}^{\epsilon_j}$ by means of insertions and/or deletions. If we cannot find such transformations, then $g \neq 1$. □

Theorem 15.4 (**[5, Theorem 5.3,p.239]**) *The word problem is solvable for any finitely generated abelian group.* (Neither a proof nor a reference is given in [5])

Theorem 15.5 (**[5, Theorem 5.4,p.240]**) *There exists a group whose word problem is not solvable.* (A reference, but not a proof is given in [5])

15.3 Solution of the word problem for the Braid group

The word problem for the braid group $\mathbf{B}_n$ is: given a braid $\beta \in \mathbf{B}_n$, is $\beta = 1$ or not? The solution consists of three steps.

Step (I) Is the braid a pure braid, that is d_i connects A_i to B_i. If not, then $\beta \neq 1$ and you are done. If yes, proceed to step (II). NOTE: β is pure if and only if its braid permutation is the identity.

Step (II) Given β a pure braid, let γ be the braid obtained from β by replacing the last string d_n by a straight line joining A_n to B_n. The set $\alpha := \beta\gamma^{-1}$. The braid α is "combed", that is, all but one of its strings is vertical. Let us write $\gamma_1 = \gamma$, $\alpha_1 = \alpha$ and repeat the process starting with $\gamma = \gamma_1$ in place of β, that is, replace the string d_{n-1} by a vertical string to get γ_2 and set $\alpha_2 = \gamma_1\gamma_2^{-1}$. Then $\beta = \alpha_1\alpha_2\gamma_2$. Continue the process until you arrive at $\beta = \alpha_1\alpha_2\cdots\alpha_{n-1}$, where each n-braid α_i is "combed".

Proposition 15.6 (**[5, Proposition 1.1,p.32]**) *Let β be a pure n-braid. Then β is the trivial braid if and only if each of the α_i in the decomposition given above is the trivial braid.*

Step (III) Determine whether or not each α_i is the trivial braid. (This is the most involved of the three steps. One shows that each α_i is an element of a free group, in which the word problem is solvable. This may be done next in this course/seminar.)

16 February 25,2005—Braid Group VI

16.1 Presentation of the Symmetric Group

17 February 28,2005—Fermat's Last Theorem I

17.1 Pythagorian Triples

17.2 Fermat's Last Theorem $n = 4$

18 March 4,2005—no class meeting

References

[1] Emil Artin, *Abhandlungen aus dem Mathematischen*, Abh. Math. Sem. Univ. Hamburg **4** (1925), 47–72.

[2] Emil Artin, *Theory of braids*, Ann. of Math (2) **48** (1947), 101–126.

[3] Percy Deift, Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach, American Mathematical Society, Providence, R. I. 2000

[4] John Derbyshire, Prime Obsession. Bernhard Riemann and the Greatest Unsolved Problem in Mathematics, Joseph Henry Press, Washington D.C. 2003

[5] Kunio Murasugi and Bohdan I. Kurpita, A Study of Braids, Kluwer Academic Publishers 1999,

[6] Norman Levinson, *A motivated account of an elementary proof of the prime number theorem*, American Mathematical Monthly **76** (1969), 225–245.

[7] Joseph Lewittes, *Quadratic irrationals and continued fractions* Number theory (New York, 1991–1995), 253–268, Springer, New York, 1996.

[8] Joseph Lewittes, *Continued fractions and quadratic irrationals*, Number theory (New York, 2003), 221–252, Springer, New York, 2004.

[9] Takashi Ono, An introduction to algebraic number theory. The University Series in Mathematics. Plenum Press, New York, 1990.

[10] Andrew M. Rockett and Peter Szüsz, Continued Fractions, World Scientific, 1992.